

# SEQUENCE SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 21:11:01 ; Search time 9362 Seconds  
(without alignments)  
11483.711 Million cell updates/sec

Title: US-09-919-831-1  
Perfect score: 2628  
Sequence: 1 cggtcacgttgggatcggtg.....ggcgacgaggggtggagctca 2628

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

Searched SEQ 1 (DNA) &  
2 (protein)

in DNA databases.

Search w/2 was a translation search.

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	2628	100.0	2628	6	AX394263	AX394263 Sequence
c	2	2628	100.0	320550	1	AP005282	AP005282 Corynebact
c	3	2628	100.0	349980	6	AX127152	AX127152 Sequence
	4	1168.4	44.5	1170	6	AX063913	AX063913 Sequence
	5	1168.4	44.5	1170	6	AX244087	AX244087 Sequence
	6	1146	43.6	1146	6	AX123145	AX123145 Sequence
	7	1146	43.6	1146	6	BD165262	BD165262 Novel pol
	8	567	21.6	567	6	AX123146	AX123146 Sequence
	9	567	21.6	567	6	BD165263	BD165263 Novel pol
	10	485	18.5	762	6	AX123144	AX123144 Sequence
	11	485	18.5	762	6	BD165261	BD165261 Novel pol
c	12	443	16.9	1041	6	AX123143	AX123143 Sequence
c	13	443	16.9	1041	6	BD165260	BD165260 Novel pol
	14	440.6	16.8	982	6	AX394270	AX394270 Sequence
	15	370	14.1	402	6	AX394269	AX394269 Sequence
c	16	353.4	13.4	834	6	AX065835	AX065835 Sequence
c	17	295.4	11.2	302070	1	AP005223	AP005223 Corynebact
c	18	240	9.1	240	6	AX123147	AX123147 Sequence
c	19	240	9.1	240	6	BD165264	BD165264 Novel pol
	20	138.2	5.3	346294	1	AP002999	AP002999 Mesorhizo
c	21	132.6	5.0	11425	1	AE001919	AE001919 Deinococc
c	22	129.6	4.9	11093	1	AE008272	AE008272 Agrobacte
	23	129.6	4.9	13715	1	AE009343	AE009343 Agrobacte
c	24	127.6	4.9	299750	1	AP005964	AP005964 Bradyrhiz
c	25	116.4	4.4	7186	1	AE009445	AE009445 Brucella
	26	113.2	4.3	11699	1	AE014493	AE014493 Brucella
	27	110.8	4.2	216750	1	BSUB0007	Z99110 Bacillus su
	28	110.2	4.2	1218	9	BT006882	BT006882 Homo sapi
	29	110.2	4.2	1218	12	BT007589	BT007589 Synthetic
	30	110.2	4.2	1326	6	AR144980	AR144980 Sequence
	31	110.2	4.2	1326	6	AX050468	AX050468 Sequence
	32	110.2	4.2	1326	9	S52784	S52784 cystathioni
	33	110.2	4.2	1831	9	BC015807	BC015807 Homo sapi

Applicant's wo

wo 02/18430  
GenBank AP005282  
EP 1108790

8/8/02

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 19:54:00 ; Search time 672 Seconds  
(without alignments)  
10556.731 Million cell updates/sec

Title: US-09-919-831-1  
Perfect score: 2628  
Sequence: 1 cggtcacgttgggatcggtg.....ggcgacgaggggtggagctca 2628

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	2628	100.0	2628	24	AAD34148	wo 02/18430
c	2	2628	100.0	349980	22	AAH68533	EP 1108790
	3	1168.4	44.5	1170	22	AAF71850	Corynebacterium gl
	4	1168.4	44.5	1170	23	AAS96112	C. glutamicum codin
	5	1146	43.6	1146	22	AAH68026	Corynebacterium gl
	6	567	21.6	567	22	AAH68027	C. glutamicum gene
	7	567	21.6	567	25	ACA00349	C glutamicum codin
	8	485	18.5	762	22	AAH68025	C glutamicum codin
c	9	443	16.9	1041	22	AAH68024	C. glutamicum deri
	10	440.6	16.8	982	24	AAD34153	C glutamicum codin
	11	370	14.1	402	24	AAD34152	C glutamicum codin
c	12	353.4	13.4	834	22	AAF71227	Corynebacterium gl
c	13	240	9.1	240	22	AAH68028	Corynebacterium gl
c	14	240	9.1	240	25	ACA00365	C glutamicum codin
c	15	222	8.4	222	25	ACA00358	C. glutamicum deri
	16	110.2	4.2	1326	22	AAC91230	C. glutamicum deri
	17	110.2	4.2	1924	23	ABV25287	Human gamma-cystat
	18	99.4	3.8	1779	22	AAS03014	Human prostate exp
	19	97.6	3.7	1733	20	AAV99909	Human diagnostic a
	20	94.6	3.6	1185	23	AAS52784	Fragment of cystat
c	21	93.6	3.6	349980	24	ABQ81847	Enterococcus faeca
	22	93.6	3.6	4403765	22	AAI99683	Bifidobacterium lo
	23	93.6	3.6	4411529	22	AAI99682	Mycobacterium tube
c	24	87.4	3.3	2944528	24	ABA03041	Mycobacterium tube
c	25	86.6	3.3	31702	24	ABQ67190	Listeria monocytog
c	26	86.6	3.3	3011208	24	ABQ69245	Listeria innocua c
	27	86.2	3.3	1302	23	AAS56046	Listeria innocua D
	28	86.2	3.3	1395	22	AAH43576	Salmonella typhi D
	29	86.2	3.3	1395	22	AAA89292	Cystathionine beta
	30	86.2	3.3	1644	21	AAC37291	Arabidopsis cystat
	31	86	3.3	78845	21	AAA81463	Arabidopsis thalia
c	32	86	3.3	349980	21	AAF21608	N. meningitidis pa
c	33	86	3.3	1437668	21	AAA81490	Neisseria meningit
	34	85.8	3.3	2411	24	ABQ70794	N. meningitidis B
	35	85.6	3.3	1209	21	AAC38726	Listeria monocytog
	36	84.8	3.2	1889	21	AAC36342	Arabidopsis thalia
	37	84.2	3.2	1161	23	AAS52693	Arabidopsis thalia
	38	84	3.2	1176	25	ABA00778	E. coli DNA for ce
	39	84	3.2	1398	25	ABA00777	N-terminal domain-
	40	84	3.2	1488	25	ABA00780	90 bp domain delet
	41	84	3.2	1692	24	ABZ14817	Transit peptide de
	42	84	3.2	1692	25	ABA00779	Arabidopsis thalia
	43	83	3.2	2365589	24	ABA90521	Full length Arabid
	44	82.4	3.1	1129	21	AAC34676	Genomic sequence o
	45	82	3.1	1092	25	ABX07272	Arabidopsis thalia
							S. pneumoniae type

# ALIGNMENTS

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 21:54:45 ; Search time 180 Seconds  
(without alignments)  
6444.194 Million cell updates/sec

Title: US-09-919-831-1  
Perfect score: 2628  
Sequence: 1 cggtcacgttgggatcggtg.....ggcgacgaggggtggagctca 2628

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	% Match	Query Length	DB ID	Description
	1	110.2	4.2	1326	3	US-09-318-448-28 Sequence 28, Appl
	2	97.2	3.7	36800	4	US-08-311-731A-139 Sequence 139, App
	3	93.6	3.6	4403765	3	US-09-103-840A-2 Sequence 2, Appli
	4	93.6	3.6	4411529	3	US-09-103-840A-1 Sequence 1, Appli
	5	86.2	3.3	1395	3	US-09-586-719-3 Sequence 3, Appli
c	6	82	3.1	11864	4	US-08-961-527-61 Sequence 61, Appl
c	7	80.2	3.1	1830121	4	US-09-557-884-1 Sequence 1, Appli
c	8	80.2	3.1	1830121	4	US-09-643-990A-1 Sequence 1, Appli
	9	77.8	3.0	1615	1	US-08-176-413-4 Sequence 4, Appli
	10	77.8	3.0	1615	4	US-09-347-878-13 Sequence 13, Appl
	11	77.8	3.0	1615	5	PCT-US94-14919-4 Sequence 4, Appli
	12	73.8	2.8	1392	4	US-09-252-991A-408 Sequence 408, App

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 22:06:10 ; Search time 916 Seconds  
(without alignments)  
9962.912 Million cell updates/sec

Title: US-09-919-831-1  
Perfect score: 2628  
Sequence: 1 cggtcacgttgggatcggtg.....ggcgacgaggggtggagctca 2628

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	% Query				DB	ID	Description
	No.	Score	Match	Length			
c	1	2628	100.0	2628	10	US-09-919-831-1	Sequence 1, Appli
	2	2628	100.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
	3	1168.4	44.5	1170	11	US-09-746-660A-79	Sequence 79, Appl
	4	1146	43.6	1146	10	US-09-738-626-3061	Sequence 3061, Ap
c	5	567	21.6	567	10	US-09-738-626-3062	Sequence 3062, Ap
	6	485	18.5	762	10	US-09-738-626-3060	Sequence 3060, Ap
	7	443	16.9	1041	10	US-09-738-626-3059	Sequence 3059, Ap
	8	440.6	16.8	982	10	US-09-919-831-8	Sequence 8, Appli
c	9	370	14.1	402	10	US-09-919-831-7	Sequence 7, Appli
	10	240	9.1	240	10	US-09-738-626-3063	Sequence 3063, Ap
	11	110.8	4.2	1173	12	US-10-369-493-41386	Sequence 41386, A
	12	100.8	3.8	1173	12	US-10-369-493-40920	Sequence 40920, A
c	13	100.6	3.8	1581	10	US-09-931-457A-67	Sequence 67, Appl
	14	99.6	3.8	1140	12	US-10-369-493-24163	Sequence 24163, A
	15	97.6	3.7	1733	10	US-09-931-457A-37	Sequence 37, Appl
	16	97	3.7	1699	10	US-09-931-457A-71	Sequence 71, Appl
c	17	95.4	3.6	1125	12	US-10-369-493-32005	Sequence 32005, A
	18	94.6	3.6	1185	9	US-09-815-242-6421	Sequence 6421, Ap
	19	94.2	3.6	1685	10	US-09-931-457A-69	Sequence 69, Appl
	20	93.6	3.6	1089	12	US-10-369-493-31668	Sequence 31668, A
c	21	90	3.4	1053	12	US-10-369-493-37724	Sequence 37724, A
	22	89.8	3.4	789	10	US-09-931-457A-33	Sequence 33, Appl
	23	89	3.4	1128	12	US-10-369-493-44368	Sequence 44368, A
	24	88.2	3.4	1152	15	US-10-156-761-3292	Sequence 3292, Ap
c	25	88.2	3.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
	26	87.4	3.3	1140	12	US-10-369-493-40221	Sequence 40221, A
	27	87.2	3.3	1143	12	US-10-369-493-32030	Sequence 32030, A
	28	86.4	3.3	1191	12	US-10-369-493-27630	Sequence 27630, A
c	29	86.2	3.3	1302	9	US-09-815-242-9683	Sequence 9683, Ap
	30	84.4	3.2	1119	12	US-10-369-493-40265	Sequence 40265, A
	31	84.2	3.2	1161	9	US-09-815-242-6330	Sequence 6330, Ap
	32	84.2	3.2	1161	12	US-10-369-493-47328	Sequence 47328, A
c	33	84	3.2	1692	10	US-09-938-842A-2622	Sequence 2622, Ap
	34	83.6	3.2	1071	12	US-10-369-493-42984	Sequence 42984, A
	35	83	3.2	1143	12	US-10-369-493-42062	Sequence 42062, A
	36	82.8	3.2	1218	12	US-10-369-493-41227	Sequence 41227, A
c	37	82.6	3.1	1116	12	US-10-369-493-35572	Sequence 35572, A
	38	82	3.1	1095	9	US-09-815-242-9554	Sequence 9554, Ap
	39	81.6	3.1	1074	12	US-10-369-493-42985	Sequence 42985, A
	40	80.8	3.1	1158	10	US-09-738-626-3491	Sequence 3491, Ap
c	41	80.2	3.1	1146	9	US-09-815-242-6865	Sequence 6865, Ap
	42	80.2	3.1	1830121	15	US-10-329-960-1	Sequence 1, Appli
	43	79.6	3.0	1122	12	US-10-369-493-32861	Sequence 32861, A
	44	79.4	3.0	1056	12	US-10-369-493-44385	Sequence 44385, A
	45	78.8	3.0	1089	12	US-10-369-493-42189	Sequence 42189, A

## ALIGNMENTS

RESULT 1  
US-09-919-831-1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 9, 2004, 21:45:20 ; Search time 4931 Seconds  
(without alignments)  
12953.187 Million cell updates/sec

Title: US-09-919-831-1  
Perfect score: 2628  
Sequence: 1 cggtcacgttgggatcggtg.....ggcgacgaggggtggagctca 2628

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*



29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	102.4	3.9	675	14	CA757379	CA757379 BR0500100
	2	100.6	3.8	2122	11	AY109661	AY109661 Zea mays
	3	99.8	3.8	563	10	BE493057	BE493057 WHE0562_H
c	4	99.6	3.8	755	12	BJ373401	BJ373401 BJ373401
	5	99.2	3.8	680	13	BW055892	BW055892 BW055892
	6	98.8	3.8	541	13	BW215645	BW215645 BW215645
	7	98.6	3.8	1137	14	CA616422	CA616422 wr1.pk009
c	8	98.2	3.7	742	12	BJ436133	BJ436133 BJ436133
c	9	98	3.7	672	12	BJ374847	BJ374847 BJ374847
c	10	98	3.7	673	12	BJ429423	BJ429423 BJ429423
c	11	98	3.7	673	12	BJ432935	BJ432935 BJ432935
c	12	98	3.7	691	12	BJ400379	BJ400379 BJ400379
c	13	98	3.7	692	12	BJ400202	BJ400202 BJ400202
c	14	98	3.7	693	12	BJ376289	BJ376289 BJ376289
c	15	98	3.7	700	12	BJ428764	BJ428764 BJ428764
c	16	98	3.7	715	12	BJ429125	BJ429125 BJ429125
c	17	98	3.7	715	12	BJ432663	BJ432663 BJ432663
c	18	98	3.7	725	12	BJ435608	BJ435608 BJ435608
c	19	98	3.7	730	12	BJ430250	BJ430250 BJ430250
c	20	98	3.7	732	12	BJ429808	BJ429808 BJ429808
c	21	98	3.7	739	12	BJ372181	BJ372181 BJ372181
c	22	98	3.7	742	12	BJ433120	BJ433120 BJ433120
c	23	98	3.7	743	12	BJ434572	BJ434572 BJ434572
c	24	98	3.7	761	12	BJ434238	BJ434238 BJ434238
c	25	98	3.7	762	12	BJ372619	BJ372619 BJ372619
c	26	98	3.7	762	12	BJ429747	BJ429747 BJ429747
c	27	98	3.7	778	12	BJ399937	BJ399937 BJ399937
c	28	98	3.7	813	12	BJ372197	BJ372197 BJ372197
c	29	96.4	3.7	721	12	BJ428519	BJ428519 BJ428519
c	30	96.4	3.7	737	12	BJ398392	BJ398392 BJ398392
c	31	95.8	3.6	701	12	BJ398232	BJ398232 BJ398232
	32	95.6	3.6	641	13	BW267916	BW267916 BW267916
	33	94.2	3.6	918	28	AZ546359	AZ546359 ENTEL37TR
c	34	94	3.6	721	12	BJ377576	BJ377576 BJ377576
c	35	93.6	3.6	659	12	BJ373114	BJ373114 BJ373114
c	36	93.6	3.6	660	12	BJ428697	BJ428697 BJ428697
c	37	93.6	3.6	660	12	BJ430241	BJ430241 BJ430241
c	38	93.6	3.6	660	12	BJ433748	BJ433748 BJ433748
c	39	93.6	3.6	667	12	BJ399583	BJ399583 BJ399583
c	40	93.6	3.6	667	12	BJ428675	BJ428675 BJ428675
c	41	93.6	3.6	668	12	BJ399430	BJ399430 BJ399430
c	42	93.6	3.6	669	12	BJ431076	BJ431076 BJ431076
c	43	93.6	3.6	685	12	BJ398580	BJ398580 BJ398580
c	44	93.6	3.6	688	12	BJ397893	BJ397893 BJ397893
c	45	93.6	3.6	692	12	BJ377815	BJ377815 BJ377815

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 10, 2004, 08:52:35 ; Search time 2373 Seconds  
(without alignments)  
3258.289 Million cell updates/sec

Title: US-09-919-831-2  
Perfect score: 947  
Sequence: 1 MEDDLAALVKALFDARTQR.....NPTTVATRYLVALDKRVPRA 189

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

Q=/cg2\_1/USPTO\_spool\_p/US09919831/runat\_09012004\_165716\_644/app\_query.fasta\_1.3  
27

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09919831\_@CGN\_1\_1\_3745\_@runat\_09012004\_165716\_644 -NCPUs=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*

11: gb\_sts:\*  
 12: gb\_sy:\*  
 13: gb\_un:\*  
 14: gb\_vi:\*  
 15: em\_ba:\*  
 16: em\_fun:\*  
 17: em\_hum:\*  
 18: em\_in:\*  
 19: em\_mu:\*  
 20: em\_om:\*  
 21: em\_or:\*  
 22: em\_ov:\*  
 23: em\_pat:\*  
 24: em\_ph:\*  
 25: em\_pl:\*  
 26: em\_ro:\*  
 27: em\_sts:\*  
 28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length	Length	Score			
	1	947	100.0	567	6	AX123146	EP 1108790	AX123146	Sequence
	2	947	100.0	567	6	BD165263	JP 2002171370	BD165263	Novel pol - Nakagawa
	3	947	100.0	2628	6	AX394263	WO 02/18430	AX394263	Sequence 7/4/02
c	4	947	100.0	320550	1	AP005282	GenBank AP005282	AP005282	Corynebact
c	5	947	100.0	349980	6	AX127152	EP1108790	AX127152	Sequence
c	6	383	40.4	329709	1	AP002997		AP002997	Mesorhizo
	7	347.5	36.7	10148	1	AE008997		AE008997	Agrobacte
	8	347.5	36.7	11534	1	AE007964		AE007964	Agrobacte
c	9	334	35.3	10998	1	AE004614		AE004614	Pseudomon
	10	323	34.1	204050	1	AL646073		AL646073	Ralstonia
	11	307.5	32.5	292200	1	SCO939129		AL939129	Streptomy
	12	297	31.4	306650	1	SME591783		AL591783	Sinorhizo
	13	289.5	30.6	300900	1	AP005939		AP005939	Bradyrhiz
	14	267	28.2	10022	1	AE005873		AE005873	Caulobact
	15	266	28.1	222605	1	AP002555		AP002555	Escherich
	16	210.5	22.2	300956	1	AE016963		AE016963	Coxiella

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 10, 2004, 08:49:14 ; Search time 258 Seconds  
(without alignments)  
1977.495 Million cell updates/sec

Title: US-09-919-831-2  
Perfect score: 947  
Sequence: 1 MEDDLSAALVKALFDARTQR.....NPTTVATRYLVALDKRVPRA 189

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

-

Q=/cgn2\_1/USPTO\_spool\_p/US09919831/runat\_09012004\_165716\_638/app\_query.fasta\_1.3  
27

-DB=N\_Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09919831\_@CGN\_1\_1\_396\_@runat\_09012004\_165716\_638 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:\*

12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:\*  
 13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:\*  
 14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:\*  
 15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:\*  
 16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:\*  
 17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:\*  
 18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:\*  
 19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:\*  
 20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:\*  
 21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:\*  
 22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:\*  
 23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:\*  
 24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:\*  
 25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%						
No.	Score	Query	Match	Length	DB	ID	Description	
	1	947	100.0	567	22	AAH68027	EP 1108790	C glutamicum codin
	2	947	100.0	567	25	ACA00349	DE 10128510	C. glutamicum deri
	3	947	100.0	2628	24	AAD34148	Wo 02/18430	Corynebacterium gl
c	4	947	100.0	349980	22	AAH68533	EP 1108790	C glutamicum codin
	5	163	17.2	17083	25	ABZ58812		S. cinnamoneus cin
	6	147	15.5	663	24	ABK78585		Bacillus clausii g
	7	146	15.4	4975	25	ABZ58814		DNA sequence of pC
c	8	129	13.6	37286	23	AAS59522		Propionibacterium
	9	127	13.4	7953	18	AAV74463		Staphylococcus aur
	10	119	12.6	240	21	AAZ91472		Sphaerotilus natan
	11	113	11.9	662	9	AAN81167		Operator from Baci
	12	112.5	11.9	552	24	ABK74142		Bacillus lichenifo
	13	112	11.8	1378	23	AAS89064		DNA encoding novel
	14	110	11.6	362	24	ABN20810		Human ORFX polynuc
c	15	108	11.4	1191	23	AAS89918		DNA encoding novel
c	16	108	11.4	1191	23	AAS93813		DNA encoding novel
c	17	106	11.2	349980	22	AAH68530		C glutamicum codin
	18	105.5	11.1	9515	15	AAQ55145		Pseudomonas aerugi
	19	105.5	11.1	9515	25	ABZ77356		Nucleotide sequenc
	20	105.5	11.1	9515	25	AAL51844		Pseudomonas aerugi
	21	103	10.9	291	25	ABX06855		S. pneumoniae type
c	22	103	10.9	5857	19	AAV52161		Streptococcus pneu
c	23	103	10.9	1830121	17	AAT42063		Haemophilus influe
	24	103	10.9	2162598	25	ABS56454		Streptococcus pneu
	25	101	10.7	240	21	AAA13364		SgrAI control prot
	26	100	10.6	540	22	AAH52849		S. epidermidis ope
	27	100	10.6	552	24	ABN91111		Staphylococcus epi
c	28	100	10.6	3364	22	AAH54269		S. epidermidis gen
	29	99.5	10.5	1799	23	AAS93668		DNA encoding novel
	30	99.5	10.5	21185	21	AAA63350		Streptomyces globi
	31	99.5	10.5	63164	21	AAA63348		Streptomyces globi
c	32	98.5	10.4	16870	20	AAX13035		Enterococcus faeca
c	33	98.5	10.4	16870	24	ABS98830		Enterococcus faeca

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 10, 2004, 11:21:50 ; Search time 79 Seconds  
(without alignments)  
1055.967 Million cell updates/sec

Title: US-09-919-831-2  
Perfect score: 947  
Sequence: 1 MEDDLAALVKALFDARTQR.....NPTTVATRYLVALDKRVPRA 189

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

Q=/cgn2\_1/USPTO\_spool\_p/US09919831/runat\_09012004\_165717\_656/app\_query.fasta\_1.3  
27

-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09919831@cgn2\_1\_1\_85@runat\_09012004\_165717\_656 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result	% Query						Description
	No.	Score	Match	Length	DB	ID	
c	1	334	35.3	732	4	US-09-252-991A-13576	Sequence 13576, A
	2	334	35.3	1098	4	US-09-252-991A-13766	Sequence 13766, A
	3	334	35.3	1731	4	US-09-252-991A-13674	Sequence 13674, A
c	4	326	34.4	1347	4	US-09-252-991A-13659	Sequence 13659, A
	5	265.5	28.0	603	4	US-09-328-352-2710	Sequence 2710, Ap
c	6	195	20.6	891	4	US-09-252-991A-13471	Sequence 13471, A
	7	171	18.1	46819	4	US-09-453-702B-72	Sequence 72, Appl
	8	157	16.6	636	4	US-09-252-991A-8938	Sequence 8938, Ap
	9	155	16.4	531	4	US-09-252-991A-8719	Sequence 8719, Ap
	10	129.5	13.7	561	4	US-09-252-991A-5411	Sequence 5411, Ap
	11	129.5	13.7	741	4	US-09-252-991A-5419	Sequence 5419, Ap
	12	129.5	13.7	879	4	US-09-252-991A-5407	Sequence 5407, Ap
c	13	129.5	13.7	1593	4	US-09-252-991A-5392	Sequence 5392, Ap
	14	121.5	12.8	783	4	US-09-252-991A-16173	Sequence 16173, A
	15	119	12.6	240	3	US-09-143-776-5	Sequence 5, Appli
	16	118.5	12.5	627	4	US-09-252-991A-7163	Sequence 7163, Ap
	17	105.5	11.1	9515	1	US-08-920-812-13	Sequence 13, Appl
	18	105.5	11.1	9515	1	US-08-920-827-13	Sequence 13, Appl
	19	105.5	11.1	9515	1	US-08-921-177-13	Sequence 13, Appl
	20	105.5	11.1	9515	1	US-08-362-577C-13	Sequence 13, Appl
	21	105.5	11.1	9515	2	US-08-920-828-13	Sequence 13, Appl
	22	103.5	10.9	582	4	US-09-252-991A-15732	Sequence 15732, A
c	23	103	10.9	5857	4	US-08-961-527-28	Sequence 28, Appl
c	24	103	10.9	1830121	4	US-09-557-884-1	Sequence 1, Appli
c	25	103	10.9	1830121	4	US-09-643-990A-1	Sequence 1, Appli
	26	101	10.7	240	3	US-09-146-675-3	Sequence 3, Appli
	27	100.5	10.6	639	4	US-09-328-352-2047	Sequence 2047, Ap
	28	100	10.6	552	4	US-09-134-001C-574	Sequence 574, App
	29	100	10.6	1062	4	US-09-252-991A-3271	Sequence 3271, Ap
	30	99	10.5	570	4	US-09-252-991A-12340	Sequence 12340, A
c	31	98	10.3	28882	4	US-08-961-527-140	Sequence 140, App
c	32	94.5	10.0	1347	4	US-09-252-991A-5399	Sequence 5399, Ap
	33	94	9.9	855	4	US-09-252-991A-867	Sequence 867, App
c	34	94	9.9	1085	4	US-08-624-447-3	Sequence 3, Appli
c	35	91	9.6	46899	1	US-08-471-119A-1	Sequence 1, Appli
c	36	89.5	9.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	37	89.5	9.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	38	88.5	9.3	1851	3	US-08-303-861-11	Sequence 11, Appl
	39	88.5	9.3	23673	4	US-09-773-816-1	Sequence 1, Appli
	40	87.5	9.2	552	4	US-09-252-991A-9619	Sequence 9619, Ap
c	41	87.5	9.2	3006	4	US-09-252-991A-9720	Sequence 9720, Ap
	42	86.5	9.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	43	86.5	9.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	44	86	9.1	360	4	US-09-252-991A-3519	Sequence 3519, Ap
	45	86	9.1	360	4	US-09-252-991A-8563	Sequence 8563, Ap

## ALIGNMENTS

RESULT 1

US-09-252-991A-13576/c

; Sequence 13576, Application US/09252991A

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 10, 2004, 12:12:10 ; Search time 280 Seconds  
(without alignments)  
2344.014 Million cell updates/sec

Title: US-09-919-831-2  
Perfect score: 947  
Sequence: 1 MEDDLAALVKALFDARTQR.....NPTTVATRYLVALDKRVPRA 189

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2276164 seqs, 1736306516 residues

Total number of hits satisfying chosen parameters: 4552328

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

-  
Q=/cgn2\_1/USPTO\_spool\_p/US09919831/runat\_09012004\_165718\_694/app\_query.fasta\_1.3  
27

-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09919831\_@CGN\_1\_1\_85\_@runat\_09012004\_165718\_694  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*



12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%					Description
	No.	Score	Query Match	Length	ID	
	1	947	100.0	567	10	US-09-738-626-3062
	2	947	100.0	2628	10	US-09-919-831-1
c	3	947	100.0	3309400	10	US-09-738-626-1
	4	171	18.1	46819	15	US-10-114-170-72
	5	162	17.1	603	15	US-10-156-761-7019
	6	162	17.1	9025608	15	US-10-156-761-1
	7	147	15.5	663	10	US-09-974-300-5876
	8	142	15.0	624	15	US-10-156-761-1633
c	9	142	15.0	9025608	15	US-10-156-761-1
	10	139	14.7	573	15	US-10-156-761-6696
	11	127	13.4	7953	8	US-08-781-986A-152
	12	112.5	11.9	552	10	US-09-974-300-1433
c	13	103	10.9	1830121	15	US-10-329-960-1
	14	99.5	10.5	63158	13	US-10-292-198-1
c	15	98.5	10.4	16870	10	US-09-070-927A-98
	16	97	10.2	402	10	US-09-974-300-3201
	17	97	10.2	504	10	US-09-974-300-5801
	18	96	10.1	12445	10	US-09-070-927A-242
	19	93	9.8	1774	12	US-10-369-493-35333
	20	92	9.7	312	10	US-09-738-626-3219
	21	92	9.7	671	13	US-10-027-632-126891
	22	92	9.7	671	14	US-10-027-632-126891
	23	92	9.7	723	15	US-10-156-761-2942
	24	92	9.7	3309400	10	US-09-738-626-1
	25	91.5	9.7	1299	15	US-10-156-761-1179
	26	91	9.6	25085	13	US-10-132-134-35
	27	91	9.6	50543	13	US-10-132-134-25
	28	90	9.5	474	15	US-10-156-761-3996
	29	90	9.5	1977	12	US-10-369-493-41620
	30	89.5	9.5	1041	15	US-10-156-761-1203
	31	89	9.4	671	13	US-10-027-632-126892
	32	89	9.4	671	14	US-10-027-632-126892
	33	89	9.4	3234	13	US-10-214-446-23
	34	88.5	9.3	14427	15	US-10-156-761-1540
	35	88	9.3	321	10	US-09-738-626-2152
	36	88	9.3	2079	15	US-10-156-761-6483
	37	88	9.3	2808	15	US-10-156-761-2021
	38	87.5	9.2	405	10	US-09-974-300-3648
	39	87.5	9.2	618	15	US-10-156-761-7117
	40	87	9.2	519	15	US-10-153-668-457

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 10, 2004, 10:53:25 ; Search time 2062 Seconds  
(without alignments)  
2227.714 Million cell updates/sec

Title: US-09-919-831-2  
Perfect score: 947  
Sequence: 1 MEDDLAALVKALFDARTQR.....NPTTVATRYLVALDKRVPRA 189

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delect 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp

-

Q=/cgn2\_1/USPTO\_spool\_p/US09919831/runat\_09012004\_165717\_650/app\_query.fasta\_1.3  
27

-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09919831\_@CGN\_1\_1\_2874\_@runat\_09012004\_165717\_650 -NCPUs=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*

```

12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			%					
No.	Score	Match	Query	Length	DB	ID	Description	
	1	281	29.7	1390	29	BZ551614	BZ551614	pacs1-60_
c	2	196	20.7	833	29	BZ570253	BZ570253	msh2_127_
	3	124	13.1	799	29	BZ559712	BZ559712	pacs2-164
c	4	111	11.7	1237	29	BZ553601	BZ553601	pacs1-60_
	5	107.5	11.4	1182	29	BZ578865	BZ578865	msh2_6040
c	6	100	10.6	763	12	BI918118	BI918118	603183036
	7	99.5	10.5	675	13	BU615857	BU615857	UI-H-DF0-
c	8	97	10.2	742	10	BE913123	BE913123	601668111
	9	96.5	10.2	574	10	BE580524	BE580524	kq41f11.y
	10	95	10.0	530	14	CA891585	CA891585	B0166G01-
c	11	95	10.0	636	12	BJ257188	BJ257188	BJ257188
	12	94.5	10.0	635	12	BP099095	BP099095	BP099095
	13	94	9.9	945	13	BU474935	BU474935	603762927
	14	94	9.9	984	29	BZ561825	BZ561825	pacs2-164
c	15	93.5	9.9	659	13	BW283547	BW283547	BW283547
	16	93	9.8	866	13	BU470075	BU470075	603760054
	17	93	9.8	867	10	BF584789	BF584789	602098546
	18	93	9.8	1099	12	BM913697	BM913697	AGENCOURT
	19	92.5	9.8	461	10	BF014939	BF014939	kq59e03.y
	20	92.5	9.8	809	13	BU744703	BU744703	CH1#004_C
	21	91.5	9.7	412	12	BM131706	BM131706	TgESTzya8
	22	91.5	9.7	413	12	BM131785	BM131785	TgESTzya8
	23	91.5	9.7	465	12	BM176679	BM176679	TgESTzya9
	24	91.5	9.7	470	10	BF014918	BF014918	kq59b09.y
	25	91.5	9.7	528	10	BG227217	BG227217	kp98g08.y
	26	91.5	9.7	541	10	BG226065	BG226065	kq08f08.y
	27	91.5	9.7	545	12	BM004036	BM004036	TgESTzya8
	28	91.5	9.7	547	13	BU576121	BU576121	TgESTzyb8
	29	91.5	9.7	642	10	BG226115	BG226115	kq09d10.y